

Ensayo de un índice conceptual para un corpus literario de textos alineados multilingües

SANTIAGO PÉREZ ISASI

Centro de Estudos Comparatistas, Universidade de Lisboa
santiagoperez@campus.ul.pt

JOSÉ RAMÓN CARRIAZO RUIZ

Universidad Nacional de Educación a Distancia-UNED
carriazo@flog.uned.es

1. INTRODUCCIÓN: EL PORTAL ANDRÉS DE POZA Y LOS ÍNDICES

El portal *Andrés de Poza* (<http://andresdepoza.com>) es un portal especializado en la edición de textos con versiones múltiples (o “textos multiversión”) dirigido por Carmen Isasi y desarrollado, en diversas etapas, con el apoyo de la Universidad de Deusto, el Gobierno Vasco y la Fundación Jesús de Gangoiti. Los textos incluidos en el corpus pertenecen a dos categorías:

- Textos que han tenido diversas ediciones o versiones a lo largo del tiempo (principalmente, textos notariales o jurídicos, como el Fuero de Alcalá de Henares, aunque también algunos textos literarios como *La libra* de Virgilio Malvezzi, en sus ediciones de Pamplona, Nápoles y Turín).
- Textos que han sido traducidos a varias lenguas. La mayor parte de los textos son literarios (por ejemplo, *Il Romulo* de Malvezzi o la *Introducción a la vida devota* de Francisco de Sales), aunque también hay ejemplos de documentos notariales multilingües, como la Fundación de la Cofradía de Pescadores de Bermeo.

El portal Andrés de Poza ofrece ediciones digitales múltiples alineadas dinámicamente. La visualización dinámica de los textos permite que el lector pueda seleccionar las versiones que quiere cotejar, hasta un máximo de cinco, que se presentan en columnas y alineados por unidad de traducción. Se ha empleado, para ello, un alineamiento previo semi-automático, llevado a cabo con el *software* de creación de memorias de traducción *Déjà Vu*, así como un etiquetado XML adaptado a las directrices de la Text Encoding Initiative (TEI-P5).

Desde el momento de su creación, el portal Andrés de Poza ha experimentado con diversos modos de enriquecimiento del texto y de visualización, que permitan al usuario-lector obtener tanta información como sea posible sobre los

SANTIAGO PÉREZ ISASI Y JOSÉ RAMÓN CARRIAZO RUIZ

«Ensayo de un índice conceptual para un corpus literario de textos alineados multilingües»
Humanidades Digitales: desafíos, logros y perspectivas de futuro. Janus, Anexo 1 (2014), pp. 347-359



Fig. 1. Página principal del Portal Andrés de Poza.

Teme d'ogn'vno il Tiranno, ed è fatale, che tema il proprio esempio, perche dal temer tutti non si escluda, in vn certo modo, il temer anche se stesso.	The Tyrant is afraid of every one, and it is fatal that he fears his owne example; for in that hee fears every one, he is not exempted from fearing, in a manner, himselfe.	Teme de cada uno el tirano, y es fatal que tema el propio exemplo, porque del temer a todos no se excluya, en un cierto modo, el temerse también a sí mismo.	Vn Tyran redoute tout; aussi est il bien iuste qu'il craigne l'effet de l'exemple qu'il donne: dans cette crainte generale de toutes choses, il semble qu'en certaine façon, la crainte de soy mesme y soit comprise.	nemen. Een Tyran vreest voor een yder, en het is een onvermydelijke nootwang, dat hy sijn eigen voorbeeld vreese, door dien van alle vreesen, in eeniger voegen niet uitgesloten word, ook sich selfs te vreesen.
Crede assicurarsi senza sangue bastoelmente dal fato, ponendola fra le Vestali, a consacrare la Virginità a gli Dii.	He thinks to fence himselfe sufficiently against fate without blood, by putting her amongst the Vestals, to consecrate her virginity to the gods.	Cree que se asegura sin sangre bastantemente del hado, poniéndola entre las vestales, y consagrando la virginidad a los dioses.	Amulus se persuade qu'en exterminant son sang et sa race, il se fortifie contre la destinee: Cela le fait resondre de mettre sa niece avec les Vestales, et de consacrer sa Virginité aux Dieux pour luy oster le moyen d'auoir lignee.	hy geloofde sonder bloedstorting sich genoegsam te versekeren wegens de hooge schikking des Noodlots, met haer onder de Verstaelsche Maegden te stellen, en haren maegdom aen de Goden te doen opofferen.
Serue alle donne, co' Tiranni, per innocenza la debolezza. Hanno eghino maggior difficultà, doue hanno minor resistenza. Non possono trouar in esse quel delitto, che fà lodare la crudeltà, o fingere in se stessi quel timore, che la fà temere.	Womens weaknesse serveth for innocencie with Tyrants; they meet with more difficulty, where they are least withstood; they cannot finde in them those faults which makes their cruelty be praised; or faine such feare in themselves as makes them be feared.	Serve a las mugeres con los tiranos la debilidad por inocencia. Tienen ellos mayor dificultad donde hallan menor resistencia. No pueden hallar en ellas aquel delito que haze alabar la crueldad, o fingir en sí aquel temor que la haze temer.	L'impuissance de ce sexe passe ordinairement pour innocence entre les Tyrans; mais les choses qu'ils redoutent le moins, sont celles qui leur font souuent plus de mal. Les femmes ne leur fournissent pas d'assez forts pretextes pour rendre leur cruauté raisonnable.	De swakheit dient de Vrouwen, in de oogen der Tyrannen, voor onnoosilheit. Sy vinden meerder moeilijckheit, waer sy minder tegenstand hebben. Sy konnen in haer sulke gebreken niet ontdecken, door welke hare wreedheit sou mogen redeneerbaer worden.

Fig. 2. Textos alineados dinámicamente.

documentos editados, así como diversas aproximaciones al texto que respondan a intereses y finalidades diversas. En este momento, además de continuar ampliando el corpus de textos y versiones en las líneas anteriormente indicadas, los esfuerzos del proyecto Andrés de Poza se centran en la creación de índices de diverso tipo que faciliten al usuario el acceso al texto y a la información que contiene. Es esta una posibilidad que venimos desarrollando desde prácticamente

el inicio del proyecto Andrés de Poza. Así, hemos implementado ya, en los textos publicados en Andrés de Poza, los siguientes tipos de índices:

- *Glosarios multilingües con y sin contexto*: Los glosarios léxicos incluidos en obras como el *Romulo* de Malvezzi ofrecen un listado multilingüe de términos significativos, que pueden visualizarse aisladamente o en su contexto amplio.
- *Índices de antropónimos y topónimos* creados automáticamente a partir del etiquetado, en obras como la *Philotea* de Francisco de Sales o *La libra* de Virgilio Malvezzi; en este último caso se han introducido diversas variantes gráficas de los topónimos ("Alsacia", "Alsace" y "Elsass", por ejemplo) para facilitar las búsquedas. En ambos casos (en el índice de antropónimos y de topónimos), es posible acceder a una lista con todas las apariciones del término con contexto, pinchando en el término deseado en cualquiera de sus manifestaciones gráficas.
- *Índices de contenidos*: En el caso de las Cartas Puebla de Vizcaya, que fueron objeto del proyecto *Foru Ondarea* (financiado por la Diputación Foral de Bizkaia. Bizkaiku Foru Aldundia), se ofrece al usuario un índice por contenidos, un tipo de índice favorecido por la estructura jurídica común a estos documentos.
- *Índices conceptuales*: El siguiente paso en esta línea de exploración, del que se ocupa precisamente esta presentación, es la creación de índices conceptuales jerárquicos y multilingües, que respondan a la estructura de un campo semántico determinado (el del léxico político, en este caso) y que permitan el acceso a los textos a partir de los conceptos.

Será necesario, antes de ofrecer una visión del estado actual de nuestras pruebas-piloto, establecer las bases teóricas y metodológicas previas en cuanto a los conceptos de "índice conceptual" y "léxico político", para después ejemplificar su aplicación al texto del *Romulo* de Virgilio Malvezzi.

2. ÍNDICES, GLOSARIOS, ONOMASIOLOGÍA Y WEB SEMÁNTICA

2.1. Planteamientos teóricos y metodológicos

La filología, en este nuevo mundo de las humanidades digitales, vuelve a ganar actualidad e interés, pues sus tareas principales (la edición de textos y la creación de instrumentos para su interpretación en sentido amplio) resultan al tiempo los objetivos básicos perseguidos por los creadores de los materiales virtuales basados en la palabra. Quien fue pionero en la Text Encoding Initiative (TEI), Christopher M. Sperberg-McQueen reclamó la figura de Mercurio como patrono de la informática y abogó así, de nuevo, por la celebración del matrimonio entre Filología y Mercurio en la era del procesamiento digital de datos:

Cristopher M. Sperberg-McQueen, Pionier der *Text Encoding Initiative* (TEI), einem Konsortium, das Standards für die digitale Textdarstellung entwickelt, nimmt Merkur als den Schutzherrn der Informatik in Anspruch und feiert so noch einmal die Hochzeit der Philologie und des Merkur in der digitalen Datenverarbeitung (Sperberg-McQueen 2001: 3-4).

Desde sus más remotos orígenes, la filología se ha encargado de la elaboración de glosas, índices, glosarios, diccionarios (sema- y onomasiológicos) y otros instrumentos cuyo objetivo es la hermenéutica o interpretación de los textos. En esa característica se basa la definición de Karl Stackmann, quien entiende la filología como ciencia de los textos antiguos y de su comprensión y cuyos instrumentos son precisamente diccionarios, gramáticas, ediciones e historias de la literatura:

Eine *zweite* Option, das philologische Tun einzufangen, liegt in der Definition von Karl Stackmann [...], der in der Philologie eine "Wissenschaft" sieht, welche "Texte der Vergangenheit verfügbar macht und ihr Verständnis erschließt". [...] Wörterbücher, Grammatiken, Editionen, Literaturgeschichten tragen die Ergebnisse in die Seminare, in die Schulen, in die Öffentlichkeit. [...] Eine pure Erschließungsphilologie führt in eine transparente, aber uninteressante Welt uniformer Metadaten, eine reine Unverfügbarkeitsphilologie in eine Welt del falschen Demut, des arkanen Raunens (Lepper, 2012: 38-39).

Tras un siglo XX que ha visto el ocaso de las filologías, el comienzo del siglo XXI, de la mano de la revolución digital, ha propiciado un resurgir del interés por los métodos y técnicas filológicas¹. Proyectos de indización semántica como Wordnet, que permiten la construcción de ontologías semánticas, posibilitan la creación de índices aplicados a corpus textuales digitales. Buscadores, corpus lematizados, bases de datos léxicos y ediciones alineadas de diccionarios (como los disponibilizados por la Fundación Rafael Lapesa) están revolucionando el mundo de las aplicaciones y herramientas informáticas para trabajar con el vocabulario desde un punto de vista diacrónico y textual.

2.2. Definición de "léxico político"

Resulta muy difícil delimitar y definir qué se entiende por "léxico político" y qué términos se incluyen, o deben incluirse, dentro de él. Esta dificultad que se observa en todas las sociedades y en cualquier momento histórico se acentúa cuando se trata de comunidades muy politizadas (Hinojo Andrés, 1997-1998: 191).

Desde la concepción foucaultiana de "arqueología del saber", "la historia del pensamiento, de los conocimientos, de la filosofía, de la literatura parece multiplicar las rupturas y buscar todos los erizamientos de la discontinuidad" (Foucault, 2009 [1969]: 15), partiendo de la revisión del documento y utilizando el concepto de discurso. A partir de estas premisas, hemos tratado de superar las dificultades a la hora de definir los conceptos propios del discurso político recurriendo a la *Begriffsgeschichte* y a la antropología lingüística; pues es indudable que existe un discurso político², para cuya elaboración debe contarse con un vocabulario característico además de unos conceptos propios y exclusivos. La

¹ "Dass der Philologiebegriff zu Beginn des 21. Jahrhunderts an Attraktivität zurückgewonnen hat, liegt nicht zuletzt an Partikularentwicklungen, am Reintegrationswillen der beteiligten Fächer und Fachquerschnitte angesichts gestiegenen Konkurrenzdrucks" (Lepper, 2012: 42).

² "Unter Diskursen werden vor allem institutionell bedingte Kommunikationsbereiche und Kommunikationsformen verstanden. Einzelne Kapitel sind hier dem politischen, dem religiösen und dem Rechtsdiskurs, dem ökonomischen und dem wissenschaftlichen Diskurs gewidmet" (Kuße, 2012: 24).

antropología política nos brinda los conceptos clave de *campo* y *arena*³, espacio conceptual donde debemos situar al autor y a los traductores del *Rómulo*, que debe redefinirse como documento complejo y a la vez instrumento del discurso político dentro de la historia de las ideas y los conceptos en la que surge. En palabras de Foucault: “la historia es lo que transforma los *documentos* en *monumentos*, y que, allí donde se trataba de reconocer por su vaciado lo que había sido, despliega una masa de elementos que hay que aislar, agrupar, hacer pertinentes, disponer en relaciones, constituir en conjuntos” (Foucault, 2009 [1969]: 16-17); la elaboración de unos índices conceptuales de las traducciones del *Rómulo* sería el *vaciado* de los documentos que componen el texto u obra, en este caso, y supondrían, precisamente referida a los conceptos políticos, el aislamiento, agrupamiento, relación, jerarquización y dotación de pertinencia de esos mismos conceptos en el ámbito del discurso político como uno de los caracteres imprescindibles de la arena como campo normativizado del juego o la competencia política en el Barroco, tal como se muestra en el *Rómulo* y en sus traducciones.

3. HACIA LOS ÍNDICES CONCEPTUALES: EL *ROMULO* DE VIRGILIO MALVEZZI

3.1. Fijación de los conceptos para el etiquetado y análisis

En una primera fase de experimentación en la creación de índices conceptuales para el portal Andrés de Poza, se realizaron tentativas de adaptación parcial de la red Wordnet, en campos semánticos más o menos autónomos que permitieran por lo tanto ser escindidos del conjunto. Sin embargo, estas tentativas fueron infructuosas debido, en primer lugar, a la escasa estructuración lógica de los campos estudiados en Wordnet; y en segundo lugar, a los problemas que supone aplicar una ontología semántica contemporánea a textos clásicos con un sistema conceptual subyacente propio de su época y culturalmente condicionado.

Para el estudio del léxico político del *Romulo* se optó, por lo tanto, por establecer en primer lugar un listado de los elementos constitutivos de la arena presentes en el texto. Se podría así empezar por destacar los individuos que entran en conflicto en la obra, es decir los nombres propios mencionados⁴, que podrían ser objeto de un índice de antropónimos como los mostrados en apartados anteriores.

Una vez excluidos estos, el índice conceptual en el que estamos trabajando se ocupa de aquellas voces cuyo significado pueda estar relacionados con los elementos básicos de todo discurso político: *apoyo, élite, estado, etnicidad, facción, género, grupos de edad, institucionalización del liderazgo, jefatura, legitimidad, matriarcado, modo de producción, nacionalismo, poder, poder/saber, reglas normativas y pragmáticas, sistema político, tribu...* (Lewellen, 2009 [2003]: 301-307). Desde el punto de vista lingüístico, se trata de léxico semántico, excluidas las palabras gramaticales: fundamentalmente verbos y sustantivos abs-

³ “Un campo político es aquel en que los grupos rivales *no* comparten reglas convenidas de antemano para regular sus conflictos. Una arena política será un área en la que los competidores aceptan las mismas reglas” (Lewellen, 2009 [2003]: 127).

⁴ Amulio, Anco, Fáustulo, Fáustulo Pastor, Lucumo, Marte, Numa, Numitor, Pan, Proca, Remo, Roma, Rómulo, Tarquino, Tibre y Tulio.

tractos, con alguna presencia de sustantivos concretos y unidades pluriverbales o sintagmáticas.

Podríamos establecer, por tanto, un primer nivel conceptual de etiquetado, de acuerdo con el agrupamiento propio de la Antropología política, tal como el presentado en la tabla siguiente:

1. Apoyo	<i>adulación, aduladores, agrado, alabança, alabar, servicio</i>
2. Élite	<i>fortuna, gozar, gozo, gracia, hijos de rey, palacio, personas de sangre real, sacerdote</i>
3. Estado	<i>arte de quitar reinos, averle quitado el reino, conquista, defender, dominio, máquina del estado, ministro, privado del reino, reino, súbdito, términos reales</i>
4. Etnicidad	<i>albanos</i>
5. Facción	<i>acciones marciales, afrenta, ambición, ardimiento, arma, combatir, debilidad, deseo, despreciar, embidia, enciende, enemigo <—> amigo, engañar, engaño, escolta, espada, fiereza, flaqueza, guerra, homicidio, huir, injuriar, inquietud, miserable, miseria, ocasión, ofensa, peligro, perder, perseguidor, pusilanimidad, rendir, resistir, resistencia, recelar, seguir, temer, temeridad, temor, tumulto, vencer, venganza, violencia, violentar, violento, victoria, vituperio, voluntario, celoso</i>
6. Género	<i>cuerpo, hombre, mujer, persona, ser humano, sexo de la hembra, varón, vestal</i>
7. Grupos de edad	<i>abuelos, acciones de los antiguos, acciones de los que passaron, antiguo, años, descendencia de Remo ('ascendencia'), edad, eternidad, favor de la fama, hechos de los presentes, juventud, origen, tiempo</i>
8. Institucionalización del liderazgo	
9. Jefatura	<i>cabeças, eminencia, furor tiránico, gloria, glorioso, grandes hechos, hechos de los príncipes, imposiciones, magestad, mandar, poder que exercen los príncipes, principado, real, regir, rey, señor, sucesor, tiranía, usurpador</i>
10. Legitimidad	<i>assegurar, beneficio, bondad, desdicha, disculpa, discurso, feliz, generoso, héroe, inferior, lícito, satisfacer, seguridad, seguro, vicios del príncipe, virtud del príncipe</i>
11. Matriarcado	<i>casa del señor ('familia del señor'), donzella, hermano, hijo, mancebo, niño, padre, sobrino</i>
12. Modo de producción	<i>acontecimientos, adquisición de bienes, caso, fabricar, interés, ocio, ocupado, pastor, pobreza, riqueza, servidumbre, valor</i>
13. Nacionalismo	<i>naciones</i>
14. Poder	<i>adoración, amor del pueblo <—> aborrecimiento del tirano, atrevido, atrevimiento, autoridad, fuerza, fuerte, obedecer, obligación, obligar, poder, poderío, tolerar, voluntad</i>

15. Poder/saber	<i>alibio, aunar, bienhechor, caridad, castigo, compadecer, compasión, condescender, condenar, convenir, cruel, crueldad, dañar, daño, dichoso, dificultad, dios, educación, elemento, elocuencia, entendimiento, galardonar, guardar la razón, hado, hombres bien afortunados <—> hombres desdichados, honestar, igual por naturaleza, impiedad, impío, infelicidad, ingenio, libertad, muerte, naturaleza, necesidad, opinión, pensamiento, piadoso, piedad, prudencia, razonamiento, reputación, sagacidad, sano consejo, severa justicia de los sacerdotes, sujeto, vergüenza, vicio, vida, virtud, vocablos de los bien afortunados, yerro</i>
16. Reglas normativas y pragmáticas	<i>armas de la elección, estrupar, ética, falsario, honra, honrar, igualar, ingratitud, injusticia, inocencia, juicio, justicia, justo, ladrón, maldad, misericordia, ninguno es superior a los otros, paciencia, sacrilegio, salvar, utilidad aiena, vil</i>
17. Sistema político	<i>arte con que está acostumbrado a vivir, consejo, pueblo</i>
18. Tribu	<i>aldea, ciudad</i>

Sin embargo, resulta necesario introducir aún, entre esta amplia clasificación de la antropología y el nivel del léxico (multilingüe) incluido en cada una de esas casillas, un nivel semántico en el que se situarían los conceptos⁵ —desprovistos, aún, de un ropaje léxico— del campo semántico político. En el ejemplo piloto que hemos desarrollado, aplicado a la edición múltiple de *Il Romulo* de Virgilio Malvezzi⁶, estamos interesados en las diversas acepciones del término “pueblo”, en relación con el campo político. Aunque existen más significados posibles, nos hemos centrado en tres de ellos:

- ‘pueblo’ como conjunto de los ciudadanos que componen un estado u otra entidad política autónoma (*demos* o *populus*);
- ‘pueblo’ como conjunto de individuos unidos por un origen y una cultura común (*ethnos*);
- ‘pueblo’ como clase social inferior opuesta a las elites dirigentes (*plebs*).

Como se puede observar, se trata de tres conceptos que, pese a poder ser encarnados —en principio— por una misma expresión léxica, se sitúan en tres casillas diferentes en la clasificación anteriormente indicada a partir de la Antropología política:

- ‘pueblo’ como *demos* se sitúa en el apartado de “estado”;
- ‘pueblo’ como *ethnos* se sitúa en el apartado de “etnicidad” o “nacionalismo”;
- ‘pueblo’ como *plebs* se sitúa en el apartado “sistema político”.

A partir de este nivel conceptual (*demos*, *ethnos*, *plebs*) es posible establecer en los textos la clasificación de las diversas expresiones léxicas en las diver-

⁵ En este contexto se utiliza el término “concepto” en un sentido utilitario y no histórico, como es el caso en el campo de la historia de los conceptos; se trata, por lo tanto, de unidades de contenido semántico como los empleados en ontologías o redes semánticas como Wordnet.

⁶ Las referencias completas de las ediciones empleadas para la edición digital multilingüe del *Romulo* de Virgilio Malvezzi se encuentran en la bibliografía final.

sas lenguas: fundamentalmente *popolo*, *pueblo*, *peuple* y *people*, pero también *gente*, *vulgo*, *nazione*, etc. Para ello, sin embargo, es necesario desarrollar un etiquetado que responda a estos diversos niveles léxico-semánticos, y que al mismo tiempo sea respetuoso con el estándar de etiquetado TEI.

3.2. Etiquetado

Dados los diversos niveles de análisis apuntados en el apartado anterior, y a la necesidad de establecer adecuadamente las interrelaciones entre las diversas versiones multilingües de la obra, se ha desarrollado un etiquetado específico basado en la estándar TEI, creado a partir del elemento <term> con un conjunto extenso de atributos que permitan definir no sólo la forma léxica lematizada de cada ítem, sino también el concepto con el que se relaciona (*demos*, *etnos* o *plebs*), y, en el caso de las versiones traducidas, la forma léxica y el concepto subyacente en el término correspondiente del original italiano. La etiqueta en su versión completa responde por lo tanto al siguiente modelo:

```
<term xml:id="t0001" xml:lang="es" key="pueblo" n="c001" sortKey="popolo"
      type="c001">
```

- @xml:id es un valor único en cada uno de los textos, que permite identificar cada *token* etiquetado y establecer la relación entre lenguas;
- @xml:lang indica mediante un código estándar la lengua del documento etiquetado;
- @key establece el lema de la palabra etiquetada;
- @n indica el concepto con el que se vincula esa aparición de la palabra (siendo "c001" *pueblo* = *demos*; "c002" *pueblo* = *plebs*, y "c003" *pueblo* = *etnos*);
- @sortKey establece la relación del término traducido con el lema del original italiano (y aparece, por lo tanto, en todas las versiones excepto el original);
- @type indica el concepto con el que se relaciona el término en el original italiano (lo que permite percibir con mayor claridad si se ha producido una modificación semántica durante el proceso de traducción).

3.3. Soluciones de visualización

Para la visualización del índice conceptual, y siempre teniendo presentes los potenciales intereses de los usuarios-lectores, se ha optado por una solución que permita percibir la jerarquización de niveles anteriormente mencionada⁷. Así, el índice conceptual desarrollado se estructura en tres columnas:

- En la primera columna, se establecen los tres conceptos diferentes anteriormente definidos: *pueblo* = *demos*, *pueblo* = *plebs*, *pueblo* = *etnos*.
- En la segunda columna se ofrecen las manifestaciones léxicas de cada uno de los tres conceptos, en el original italiano.
- En la tercera columna se ofrecen las traducciones de esas manifestaciones léxicas del original, en las restantes lenguas, independientemente de que

⁷ El desarrollo informático necesario para la visualización del índice ha sido realizado por la empresa Bitext.

mantengan el sentido del original o lo sustituyan por otro (más adelante volveremos sobre este aspecto).

Conjunto de gentes (19)	gente (4)	en: people (3)
		es: gente (4)
		fr: people (2)
	popolo (15)	en: people (14)
		en: common people (2)
		es: pueblo (15)
		es: hombre (1)
		fr: people (12)
		fr: homme (1)
		fr: nation (1)

Fig. 3. Índice conceptual para el concepto ‘demos’.

El usuario-lector puede, a continuación, ver todas las apariciones de determinada expresión léxica en una de las lenguas (pinchando encima de cualquiera de las expresiones que aparecen en la columna segunda o tercera), o bien, si lo prefiere, obtener una visión de conjunto de todos los casos relativos a uno de los conceptos en todas sus expresiones léxicas multilingües, clicando para ello en el concepto que aparece en la primera columna.

Venne sospitione al popolo , che i Senatori (quali hauerua spogliati dell'autorità) l'haussero amazzato.	Le peuple soupçonna que les Sénateurs dégradés par Romulus de l'autorité qu'il leur auoit luy mesme donnée luy auoient fait oster la vie.	Sospechó el pueblo que los senadores, a quien avia quitado la autoridad, le avian muerto.	The people did suspect that the Senators (whom he had deprived of their authority) had slaine him.
Tumultua il popolo . Alza il bollore, ma non lo spande fuori del vaso: Si mostra pronto a seguire chi voglia vendicare. Vn Senatore, che in quel punto si fosse fatto capo del popolo , si faceua, non hà dubbio, Signore della Città.	La mort de Romulus esmeut le peuple , il boult de tumulte, mais il ne sort pas pour tant pardessus les bords du vaisseau: Il ne demande qu'à remuer les mains, et disent qu'ils sont tous prests à sulure celuy qui voudra la vanger. Le beau Jeu qu'il y auoit alors pour vn ambitieux! Vn Sénateur qui se fust à cettè là déclaré chef du peuple , se fust sans doute fait chef et maistrè de la Cité.	Alborótase el pueblo , hierbe, mas no vierte fuera del vaso el herbor: muéstrase pronto a seguir al que quisiere vengança. Un senador que en aquella ocasión se huviera hecho cabeça del pueblo , se huviera hecho, sin duda, cabeça de la ciudad.	The people make a tumult; their choler boyles apace, but leapes not out of the vessell; they shew themselves readie to follow any who will revenge him; some Senator, who at that time had made himselfe the peoples Leader, might doubtlessly have made himselfe Master of the Citie.
Giulio Proculo vi accorre, raccontando hauer veduto Romulo ascendere al Cielo, e che comandaua di voler essere chiamato Dio Quirino. Il popolo lo crede; l'acqueta, ed in vece di vendicarlo, si dà a sacrificargli.	Iulius Proculus suruint là dessus qui appaisa toute la sedition, jurant et protestant qu'il auoit veu Romulus enleuer dedans le Ciel, et qu'il luy auoit commandé de dire à son peuple qu'il honorast de sormais sa memoire du nom de Quirin. Le peuple adiousta foy au rapport de Proculus, parce qu'il estoit estimé fort homme de bien: De façon qu'au lieu de rechercher la vengeance de la mort de Romulus , ils luy volüent des sacrifices comme à vn Dieu.	Julio Próculo los socorrió afirmando que avia visto subir al cielo a Rómulo, y que mandava que le llamassen Dios Quirino. El pueblo lo cree, se quieta y en lugar de vengarle le sacrifica.	Julius Proculus came in, affirming that he saw Romulus ascend up into heaven, and that his will was to be called the god Quirinus; the people believe him, appease the tumult, and in stead of revenging him, prepare to offer sacrifice unto him.

Fig. 4. Expresiones léxicas multilingües del concepto ‘demos’.

Mediante esta comparación de las diversas traducciones de un determinado concepto en el original italiano, es posible analizar con especial claridad las elecciones verbales de los traductores, o las diferencias entre las diversas lenguas en su estructuración del campo semántico y de sus representaciones léxicas. Así, por ejemplo, se observa que Henry Carey, traductor de la versión inglesa, enfatiza sistemáticamente el concepto *pueblo* = *plebs* mediante la lexía compleja *common people*.

Gli stati, che dormono quieti, perche sono amici de'vicini, hanno gran ventura, se incontrano in qualche occasione di sdegno; e gli'huomini auertiti, in così fatti casi la cercano, perchè il popolo non si lascia persuadere se non quello che vede. Egli giudica coll'occhio, non coll'intelletto; ne vi hà argomento presso di lui basteuole, per espugnare l'apparenza. Il tenere amicitia de'vicini è ottimo. Sopra di quella fondare la sicurezza dello stato è pessimo. Stanno bene, per amici, se si considerano per nemici, accò debbano amare, e non possono offendere. L'altezza di quell'edifitio, che piace, quando altri crede, che gl'habbia a seruire d'habitatione, displace, se vi si aspetta il precipitio.	Les Etats qui dorment en repos, à cause qu'ils sont amis de leurs voisins sont fort heureux quand ils rencontrent quelque suiet de courroux: Les hommes bien aduisez le cherchent tousiours en ces temps là; parce que le peuple ne peut lamais apprehender que ce qu'il voit present deuant soy; Il luge avec la veuë et non pas avec l'entendement; Il n'y a point d'argument assez fort qui luy puisse contester l'apparence. Il est bon d'auoir amitié et alliance avec l'Estat voisin, mais c'est faire mal de fonder la seureté du sien là dessus: nos voisins sont bons pour amis, quand nous les considerons comme s'ils estoient ennemis, afin qu'ils aiment seulement, et qu'il ne puissent offenser: La grandeur et hauteur d'un edifice est tres-agreable à celuy qui croit d'en faire son habitation, mais quand il preuolt que cette hauteur là doit estre son precipice, elle luy est odieuse.	Los estados que duermen quietos, porque son amigos de los vezinos, tienen gran dicha si encunturan en alguna ocasión de enojo; y los hombres advertidos en semejantes casos la buscan, porque el pueblo no se dexa de persuadir sino lo que ve. Él juzga con la vista, no con el entendimiento; ni ay argumento eficaz para él que le contraste la apariencia. El tener amistad con los vezinos es bueno. Sobre aquélla fundar la seguridad del estado es malo. Son buenos para amigos si se consideran por enemigos, para que deban amar, y no puedan ofender. La alteza de aquel edificio que agrada quando uno cree que le ha de servir de habitación, le aborrece quando le considera como precipicio.	States that enjoy rest and quiet, because they are in league and friendship with all their Neighbours, have great good fortune if they meet with any occasion of offence; and wise men in such like occasions seeke after them; for the common people will not be perswaded to more than what they see, they judge by the eye, not the understanding, nor is there any argument with them of force enough to gain-say appearance. To keepe friendship with neighbours is exceeding good; but hereupon to build the securitie of a State, is passing bad: they are well held for friends, if they bee considered as enemies; that so they may bee bound to love, and not able to hurt; the height of that building, which is delightfull and pleasing, when one thinkes to enjoy it for an habitation; is displeasing, if he looke it shall fall upon his head.
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Fig. 5. Ejemplo de selección léxica para 'pueblo = plebs'.

Por otro lado, resulta también de vital importancia indicar, en la visualización que se ofrece al lector, aquellos casos en los que se ha producido una alteración de conceptos en el proceso de traducción, mediante un código de colores: el azul indica que existe identidad entre los conceptos en las diversas lenguas; el color verde indica que se ha producido una sustitución de un concepto por otro. En el ejemplo que se muestra a continuación, el original italiano “il popolo”, que hace referencia muy probablemente al conjunto de los ciudadanos (*demos*), es sustituido por “common people” (*plebs*) en la traducción inglesa.

Hà il Principe vn non sò che più dell'huomo, nella maestà del volto, ne gl'Angioli, che lo difendono, nelle stelle, che gl'influiscono. Alcuni gli diedero nome d'Heroe, la verità la chiamò Dio, e gli gentili, non sarento trauati dal retto, se equiuocando dalla simiglianza all'essenza, al nome di Dio non haussero agglittà l'adoratione. Il popolo , perchè lo crede maggiore dell'huomo, si marauiglia se lo vede uguale, si scandalessa se lo conosce inferiore. Non deeno i Principi lasciarsi misurare. Cimentarsi al paragone senza sicurezza di vincere, è sicurezza di perdere. Vn non sò che di più, che da gl'altri si desidera, da chi hà un non sò che di più de gl'altri.	Les Princes ont tousjours quelques graces particulieres par dessus les autres hommes, soit en la maesté de l'aspect, soit que les Anges les defendent, ou bien en ce que les estoilles versent sur eux leurs benignes influences. Aucuns les ont nommez Heros, autres diuins, et les Gentils n'eussent pas excédé ce qui est licite, si prenant la ressemblance pour l'essence, ils n'y eussent adiousté l'adoration. Les hommes s'esmerueillent de les voir imperieux comme les Dieux, et se scandalisent quand ils les voyent inferieurs: Les Princes ne se dolent point laisser mesurer: se raualer à la comparaisol sans assurance de vaincre, on se doit asseurer de se perdre: Il est necessaire qu'il y ait vn le ne sçay quoy plus qu'aux autres, en celuy qui à vn ie ne sçay quoy par dessus les autres.	Tiene el príncipe un no sé qué más que hombre en la magestad del semblante, en los ángeles que le defienden, en las estrellas que le influyen. Algunos le dieron nombre de héroe, la verdad lo llamó dios, y los gentiles no hubieran excedido de lo lícito si, equivocando la semejanza a la esencia, al nombre de dios no hubieran juntado la adoración. El hombre , porque le ve mayor que el hombre , se maravilla si le ve igual, se escandaliza si le conoce inferior. No deben los príncipes dexarse medir. Baxarse a la comparación sin seguridad de vencer, es seguridad de perderse. Un no sé qué más que en los otros se desea en quien tiene un no sé qué más que los otros.	Princes have somewhat more than other men in the majestie of their countenance, in their tutelary Angels, and in the influences of the Starres. Some call a Prince an Heroe, Truth it selfe calls him a god, nor had the Gentiles bene much out, had they not by equivocating from the similitude to the essence, added to the Name of God adoration; the common people , because they thinke him more than man, wonder if he be but equall to other men, are offended if hee be inferior: Princes ought not suffer themselves bee measured; who suffers competition, not sure to win, is sure to lose; somewhat of more than from others is required from him that hath somewhat of more than others.
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Fig. 6. Sustitución de conceptos en la traducción.

Naturalmente, serían necesarias, para extraer conclusiones definitivas respecto a la configuración de los campos semánticos implicados, investigaciones más exhaustivas que tuviesen en cuenta los condicionantes lingüísticos, culturales y políticos de cada traducción, empleando para ello, entre otras fuentes documentales imprescindibles, los diccionarios históricos correspondientes a cada una de las lenguas.

4. CONCLUSIONES: POSIBLES UTILIZACIONES FUTURAS DE ESTOS ÍNDICES

Lo que hemos mostrado en las páginas anteriores es sólo una prueba piloto de creación de un índice conceptual, muy limitada por lo tanto en su alcance y pretensiones, con un reducido núcleo de conceptos interrelacionados pertenecientes al ámbito político. Con todo, y a pesar de lo limitado de la muestra escogida, la experimentación en el establecimiento de un conjunto de conceptos, su etiquetado y visualización han permitido extraer conclusiones que serán de gran utilidad para futuras experiencias en esta misma línea.

Así, la ampliación de esta experiencia piloto a un corpus más amplio, y al conjunto del léxico político europeo del siglo XVII, exigiría la creación de una red de conceptos o de una ontología propia, dado que, como hemos visto, las ontologías genéricas actualmente existentes (como WordNet) tienen obvias carencias cuando se aplican a campos semánticos que quizá no estén todavía suficientemente estandarizados, y en especial cuando se aplican a textos históricos y no contemporáneos. Es este un proceso que, como ya se ha apuntado, exigiría un riguroso trabajo filológico, que combinase una aproximación desde la antropología política, con una investigación lexicográfica multilingüe y contrastiva.

En cualquier caso, un desarrollo como el apuntado aquí, en particular si se extendiese a todo el corpus Andrés de Poza, o incluso a un corpus más amplio (sincrónico o diacrónico) permitiría realizar estudios en áreas diversas, tales como los estudios de traducción; los estudios de lexicografía; la historia de los conceptos o la antropología política. La aplicación de técnicas informáticas de tratamiento y enriquecimiento de textos se alía así, una vez más, con los objetivos de la filología en su sentido más amplio y más fructífero.

BIBLIOGRAFÍA

- Foru Ondarea, <<http://foruondarea.net/>> [10/11/2013].
- Foucault, Michel, *La arqueología del saber*, Madrid, Siglo XXI, 2009 [1969].
- Fundación Rafael Lapesa, <<http://www.frl.es/Paginas/default.aspx>> [10/11/2013].
- Hinojo Andrés, Gregorio, "El léxico político romano: *speciosa uerba*", *Voces*, 8-9 (1997-1998), pp. 191-206.
- Kuße, Holger, *Kulturwissenschaftliche Linguistik. Eine Einführung*, Göttingen, Vandenhoeck & Ruprecht, 2012.
- Lepper, Marcel, *Philologie zur Einführung*, Hamburg, Junius Verlag, 2012.
- Lewellen, Ted C., *Introducción a la Antropología Política*, Barcelona, Edicions Bellaterra, 2009 [2003].
- Malvezzi, Virgilio, *II Romulo*, Bologna, Clemente Ferroni, 1629.

Malvezzi, Virgilio, *Romulus and Tarquin*, traducción de Henry Carey, Londres, J. Haviland, 1637.

Malvezzi, Virgilio, *Le Romulus du marquis de Maluezzi, avec ses considérations politiques et morales sur sa vie*, traductor anónimo, Paris, Paris J. LeBouc, 1645.

Malvezzi, Virgilio, *El Rómulo*, traducción de Francisco de Quevedo, edición crítica de Carmen Isasi, Bilbao, Universidad de Deusto, 1993.

MultiWordNet, <<http://multiwordnet.fbk.eu/online/multiwordnet.php>> [10/11/2013].

Portal Andrés de Poza, <<http://www.andresdepoza.com>>.

Sperberg-McQueen, Christopher M., "Die Hochzeit der Philologie und des Merkur: philologische Datenverarbeitung", en Stephan Moser et al. (eds.), *Maschinelle Verarbeitung altdeutscher Texte. Beiträge zum 5. Internationalen Symposium 1997*, Tübingen, Niemeyer, 2001, pp. 3-22.

Text Encoding Initiative, <<http://www.tei-c.org>> [10/11/2013].

Wordnet Multilingual Central Repository Interface, <<http://adimen.si.ehu.es/cgi-bin/wei/public/wei.consult.perl>> [10/11/2013].



RESUMEN

En los progresos de la edición digital de textos literarios, la elaboración de índices, tesauros y glosarios ofrece todavía un amplio espacio de exploración. En el marco del proyecto Andrés de Poza (<<http://andresdepoza.com>>), dedicado a la edición de textos literarios europeos de los siglos XVI y XVII en versiones múltiples (original y varias traducciones) se ha trabajado ya en la preparación de índices de diverso tipo. En el momento actual, se ha considerado la posibilidad de integrar índices temáticos, sustentados en los propios índices de materias de los testimonios editados, o bien de nueva creación. En cuanto a la resolución técnica, contamos en este sentido con la experiencia obtenida en los índices temáticos de las *Cartas Puebla de Bizkaia* (<<http://foruondarea.net>>), elaborados a partir de una marcación XML.

Este ensayo presenta las decisiones tomadas a lo largo del proceso de creación del índice conceptual multilingüe a través de un ensayo realizado con vocabulario del campo político (un ensayo limitado pero metodológicamente exportable a otros términos del mismo campo, así como a otros campos). Finalmente, se intenta también mostrar con una breve ejemplificación las potenciales aplicaciones futuras de este tipo de índices: estudios de léxico histórico, estudios de traducción, estudios culturales comparados y sobre todo historia de los conceptos –en la línea de la antropología política–, además de su propio interés como herramienta de acceso a los textos multilingües editados.

Palabras clave: Índices, marcación, edición, textos alineados, léxico político.

ABSTRACT

In the last developments of literary texts digital editing, the making of indexes, thesauri and glossaries still offers an open field for researching. In the frame of the project "Andres de Poza" (<<http://andresdepoza.com>>), devoted to European sixteenth and seventeenth centuries literary texts editions in multiple versions (original and translations to several languages), we have already worked on the preparation of indexes of various kinds. At the moment, we have considered the possibility of integrating thematic indexes for the edited or newly created textual witnesses, supported by their own tables of contents. As for the technical resolution, we count on the experience obtained in the subject indexes of the *Cartas Puebla de Bizkaia* (<<http://foruondarea.net>>), elaborated with an XML marking code.

This paper presents the decisions made throughout the process of creating multilingual indexes of the vocabulary from the political lexicon (one methodologically limited trial but exportable to other terms in the same semantic field, as well as to other fields or lexicons). Finally, we show also potential future applications of this type of indexes: diachronic studies of the vocabulary, translation studies and traductology, comparative cultural studies and especially history of concepts –as meant in political anthropology–, as well as a tool for accessing multilingual edited texts and corpuses.

Keywords: Indexes, marking, editing, text alignment, political lexicon.

